

DISTRIBUTED COMPRESSED SENSING: SPARSITY MODELS AND RECONSTRUCTION ALGORITHMS USING ANNIHILATING FILTER

Ali Hormati, Martin Vetterli

AudioVisual Communications Laboratory
Ecole Polytechnique Federale de Lausanne
1015 Lausanne, Switzerland
{ali.hormati, martin.vetterli}@epfl.ch

ABSTRACT

Consider a scenario where a distributed signal is sparse and is acquired by various sensors that see different versions. Thus, we have a set of sparse signals with both some common parts, and some variations. The question is how to acquire such signals and how to reconstruct them perfectly (noiseless case) or approximately (noisy case). We propose an extension of the annihilating filter method [3] to this distributed scenario. We model the inter-relation between the sparse signals by introducing three joint sparse models. For each model, we propose sensing and reconstruction algorithms that reduce the number of measurements below the limit for the single sensor scenario and results in power and bandwidth reduction in the system. In the noiseless scenario, we are close to the minimum number of measurements possible for the perfect reconstruction while by taking more measurements, we introduce redundancy in the system to effectively mitigate the noise. Simulation results justify the applicability of the approach.

Index Terms— Compressed sensing, Sparse signals, Annihilating filter, Vandermonde matrix, Finite rate of innovation sampling

1. INTRODUCTION

Results on compressed sensing suggest that a relatively small set of measurements taken as the inner products between a signal and random measurement vectors can well represent a source that is sparse in some fixed basis [1], [2]. In [3], we proposed a sensing and recovery mechanism based on Vandermonde matrices and annihilating filters which is able to reconstruct k -sparse signals of dimension $n \gg k$ using just $2k$ measurements with reconstruction complexity of $O(k^2)$. This idea is in connection with the work by Vetterli et al. on sampling methods for signals with finite rate of innovation [4].

In this paper we extend our results in [3] to the distributed sensing scenario discussed in [5]. We propose new sensing and recovery mechanisms to solve the distributed sensing problem. Assume that there are N sensors in an area measuring a phenomenon both in space and time. A base station receives all the measurements and runs an algorithm to jointly decode the signals of the sensors and reconstruct the phenomenon at the sensor positions, see Figure 1. The assumption of the compressibility of the signals lets us assume that each signal is sparse in some basis. Moreover, since the sensors presumably observe related phenomena, there should exist some inter-relation between them. One should be able to make use of this inter-relation to compress the whole ensemble of signals more efficiently.

This work was supported in part by the Swiss National Science Foundation under grant 200020-103729.

Since the sensors take the measurements independently from each other, there is no need for any communication between the sensors except the routing of the measurements to the base station. In addition, the joint decoding which is much more power consuming than sensing is done on the base station and sensors do not take part on it at all. These two facts besides the compression of the joint sparse signals by the distributed sensing scenario result in less power and bandwidth consumption by the sensors which are valuable resources in a low-power sensor network.

As an example, think of a microphone network recording a sound field at several points in space. The time series acquired at any of the microphones might be sparsely represented in some Fourier basis. In addition, the group of signals of all the microphones might show a high correlation in the form of having some frequencies in common because they listen to the same source.

Assume that the signals of the N sensors are k -sparse in some orthonormal basis like Fourier or wavelet. If there exists no inter-relation between these signals, the results in [3] suggest that one needs $2kN$ measurements in total to fully represent the signals at the base station. Following the same modeling approach as in [5], in Section 2 we define three different Joint Sparse Models (JSM) to model different types of connections between the signals. In Section 3, we propose sensing and reconstruction algorithms which can exploit the JSM model and reduce the number of measurements below the limit for one sensor scenario. In the noisy case, we denoise the measurements by the Cadzow's algorithm as described in [6] and [7] before applying the annihilating filter for reconstruction. The denoising algorithm will be adapted to the distributed scenario and will benefit from joint sparse models to better mitigate the noise. This will be followed by simulation results in Section 4 which evaluate the performance of the algorithms numerically.

2. JOINT SPARSE MODELS

In this section, we define the notion of an ensemble of signals being jointly sparse and model the inter-relation between the signals of the sensors under three joint sparse models.

2.1. Sparse Common Support: JSM-1

In this model, all N signals share a common support set but with different coefficients. The signal of sensor q is

$$x_q = \Psi \theta_q, \quad (1)$$

where the indexes of the non-zero components of θ_q are in a set Υ of cardinality k_c . The set Υ is the same for all the sensors.

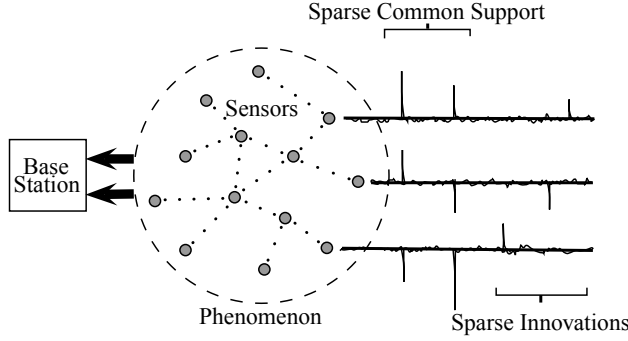


Fig. 1. Distributed sensing scenario. A collection of sensors measure a phenomenon and send their measurements to the base station in which they are jointly decoded. There exists some inter-relation between the signals.

A practical situation well-modeled by JSM-1 is where a set of acoustic sensors acquire replicas of a set of Fourier-sparse signals with different amplitudes and phases caused by attenuations and multipath effects. In applications such as acoustic localization and array processing algorithms, it is necessary to recover all of the signals of the sensors at the base station for further processing.

2.2. Sparse Common Support + Sparse Innovations: JSM-2

This model extends the JSM-1 so that there exists specific sparse innovation components in the signals. The support set and non-zero coefficients of the innovation parts are different among the sensors. Note that the common component is common only in the support set and the coefficients are different from one sensor to the other.

Assume that the common support set has cardinality k_c and k_i is the cardinality of the sparse innovation part. We assume the same cardinality for the supports among the sensors. The signal of sensor q can be represented as

$$x_q = s_q^c + s_q^i = \Psi(\theta_q^c + \theta_q^i), \quad (2)$$

where the vector θ_q^c has nonzero indexes in a set $\Upsilon \in \{1, 2, \dots, n\}$, common between the sensors, in which the cardinality of Υ is k_c and θ_q^i represents the innovation part coefficients vector with the nonzero support cardinality k_i .

A practical situation modeled by JSM-2 is where a group of sensors measure a physical phenomenon such as humidity in a region. There is a global effect of the physical quantity over the whole region which can be captured by the common part. The local physical effect at the location of each sensor is reflected in the innovation part.

2.3. Full Common Component + Sparse Innovations: JSM-3

In the third joint sparse model, there is a non-sparse common signal among the sensors plus a sparse innovation part which is specific to each one. Let the signal of sensor q be

$$x_q = x_c + s_q^i = x_c + \Psi\theta_q^i, \quad (3)$$

where θ_q^i represents the innovation part coefficients with the nonzero support cardinality k_i . The innovation part of each sensor can have its specific support set and coefficients different from the other sensors.

One example of this sparsity model can be in the compression of data such as video where each video frame is not sparse but the differences between the video frames may be sparse in some sparsity basis. It is possible to encode the frames independently and jointly recover them at the decoder.

3. SENSING AND RECOVERY MECHANISMS

The sensing mechanism at the sensors is based on the single sensor scenario in [3]. In order to take m measurements at each sensor, the measurement matrix Φ which is the same at all the sensors is represented as

$$\Phi_{m \times n} = \Gamma_{m \times n} \cdot \tilde{\Psi}_{n \times n}, \quad (4)$$

where Γ is any Vandermonde matrix and the rows of $\tilde{\Psi}$ have the orthogonal (biorthogonal) relationship with the columns of the sparsity inducing basis Ψ , $\tilde{\Psi}\Psi = \Psi\tilde{\Psi} = I$. A good choice for the matrix Γ is m consecutive rows of the n by n DFT matrix. For simplicity, one can assume that $\Psi = I$ so that the signals are sparse in the discrete dirac basis. At sensor q , the measurement vector y_q is computed as

$$y_q = \Phi x_q, \quad (5)$$

and this measurement along with the measurements of the other sensors are sent to the base station for joint recovery.

3.1. JSM-1 Recovery

3.1.1. Algorithm 1: Concentrated Common Support Recovery

In this strategy, one takes $2k_c$ measurements at sensor 1 and k_c measurements from the remaining $N - 1$ sensors. The recovery mechanism in the base station consists of first setting up an annihilating filter of length k_c of the form

$$H(z) = \prod_{i=1}^{k_c} (1 - u_i z^{-1}) = \sum_{\ell=0}^{k_c} H[\ell] z^{-\ell}. \quad (6)$$

Using the $2k_c$ measurements taken at sensor 1, the annihilating filter roots gives the common support set and the coefficients of the signal of sensor 1. By using the true support, the nonzero coefficients of the signals of the remaining $N - 1$ sensors are computed by solving a linear system of k_c equations in k_c unknowns. Since there are k_c unknown positions and Nk_c unknown coefficients in the JSM-1 model, the number of unknowns to be determined by the measurements is $k_c(N + 1)$. The total number of measurements in this strategy is $2k_c + (N - 1)k_c = k_c(N + 1)$ which is equal to the number of unknowns. On the other hand, the average number of measurements per sensor is equal to $k_c(1 + \frac{1}{N})$ which tends to k_c as the number of sensors increases. This situation is the same as the case where a genie gives the unknown common support set in advance and the measurements are used to determine the unknown coefficients. In the noisy case, we increase the number of measurements at the sensors to add redundancy to the system. The iterative Cadzow's denoising algorithm as described in [6] is then used to denoise the measurements before applying the annihilating filter.

3.1.2. Algorithm 2: Distributed Common Support Recovery

In this strategy, one distributes the task of finding the common support set among k_c sensors. Since k_c measurements are needed in each sensor for computing its unknown coefficients, there should be at least $k_c + 1$ measurements at the sensors which want to take part

in the common index set estimation. Surprisingly, by taking $k_c + 1$ measurements, it is possible to write one equation for the annihilating filter of length k_c . The annihilating filter is the same for all the signals because the signals share the same support; the same filter will annihilate the signals of all the sensors.

Since one needs k_c equations to find the annihilating filter, there should be at least k_c sensors each giving $k_c + 1$ measurements. The measurements of the remaining $N - k_c$ sensors can be of size k_c just to compute their unknown coefficients. It is also possible to take more than $k_c + 1$ measurements at any sensor and thus reduce the number of sensors which contribute to the common support recovery. The total number of measurements is $k_c(N + 1)$ which is equal to the number of unknowns, similar to strategy 1.

In the noisy setup, the measurements matrix should be of rank k_c and block Toeplitz where each block corresponds to one of the sensors. The Cadzow's denoising method finds the closest (in Frobenius norm) measurements matrix with these properties to the noisy one. After denoising, the annihilating filter will find the unknown support set. The unknown coefficients are found by the method of least squares.

3.2. JSM-2 Recovery

3.2.1. Emphasized Common Support Recovery

Although JSM-2 seems to be a simple extension of JSM-1, the reconstruction strategy is considerably different. Note that in JSM-2, there is not any single annihilating filter with minimum length that can annihilate all the signals of the sensors because of the specific innovative support set added to each signal by the innovation part.

Our approach in the recovery process is first finding the common support set by considering the innovative part as noise in the measurements. Assume that one takes $k_c + 2k_i$ measurements at each of the N sensors plus k_c more measurements distributed among them. In this case the total number of measurements is equal to the total number of unknowns (which is equal to the sum of k_c unknowns for common positions and $N(k_c + 2k_i)$ unknowns for the coefficients). The recovery process starts by building the annihilating filter of degree k_c . It is possible to write a total of $2Nk_i + k_c$ equations for the annihilating filter. This is equivalent to an oversampling factor of $\frac{2Nk_i}{k_c} + 2$ which increases by having more sensors. We solve for the common support set by looking for an annihilating filter of length $k_c + 1$ and hope that the oversampling factor just mentioned is able to mitigate the effect of the innovation part. Then, by removing the common part from the measurements of each sensor, one can easily find the innovation parts in the classical way.

The measurements matrix in the common support set estimation phase should be of rank k_c and block Toeplitz. In the innovation estimation phase, after removing the effect of the common part, the resulting measurements matrix for each sensor should be Toeplitz and of rank k_i . One can use the Cadzow's method to denoise the measurements in both phases before applying the annihilating filter.

3.3. JSM-3 Recovery

3.3.1. Algorithm 1: Common Signal Estimation Recovery

This recovery method tries to exploit the existence of a common part in the signals to first get a good estimate of the effect of the common signal on the measurements of the sensors. Then by removing this effect, the whole system goes back to the sparse scenario where one can find the innovation parts in the standard way.

Lets setup the strategy. Take the whole signal of sensor number 1 (which is of length n) and $2k_i$ measurements from the rest and send them all to the base station which results in $n + 2(N - 1)k_i$ measurements in total. In the base station, the averaged measurement vector y_a is equal to

$$\begin{aligned} y_a &= \frac{y_1 + y_2 + \dots + y_N}{N} \\ &= \Phi(x_c + \frac{s_1^i + s_2^i + \dots + s_N^i}{N}), \end{aligned} \quad (7)$$

where Φ is the measurement matrix of size $2k_i \times n$. So by increasing N , the number of sensors, one can assume that $y_a \simeq \Phi x_c$; it is possible to find the effect of the common part x_c on the measurements received from the sensors. The next step is to subtract y_a from the measurements received and find the innovation part of each of the sensors in the classical manner from the new measurements. Since the whole signal x_1 of sensor 1 is available, one can find x_c from the whole signal received from sensor 1. Having access to the common and innovative parts of the signals of the sensors, one can reconstruct all the signals at the base station.

3.3.2. Algorithm 2: Differential Innovation Recovery

In this strategy, one takes more measurements compared to the algorithm 1 but this enables the algorithm to recover the signals of the sensors perfectly. One sends the whole signal of sensor 1 and takes $4k_i$ measurements from the rest. At the receiver, one builds the equivalent measurement vector of sensor 1 of length $4k_i$ from its whole received signal with the same measurement matrix of the other sensors. Build the new measurement vectors as

$$\begin{aligned} \hat{y}_2 &= y_2 - y_1 = \Phi(s_2^i - s_1^i) \\ \hat{y}_3 &= y_3 - y_1 = \Phi(s_3^i - s_1^i) \\ &\vdots \\ \hat{y}_N &= y_N - y_1 = \Phi(s_N^i - s_1^i). \end{aligned} \quad (8)$$

Since the support of the difference of the innovation parts on the right hand sides of these equations are at most $2k_i$, it is possible to find them using $4k_i$ measurements. Once found, the signals of the sensors from 2 to N can be found by adding them to the whole signal of sensor 1 (which is $x_c + s_1^i$). In this process, the total number of measurements is equal to $n + 4(N - 1)k_s$ which is more than the number of unknowns approximately by a factor of 2 for large N but still much smaller than the naive case which needs nN measurements in total.

4. SIMULATION RESULTS

In this section we provide simulation results which demonstrates the performance of the proposed sensing and recovery mechanisms in three different joint sparse models discussed in the paper.

For the JSM-1 scenario, we consider signals of length $n = 64$ with $k_c = 4$ nonzero coefficients and number of measurements equal to $m = 2k_c = 8$. The positions of the non-zero coefficients are chosen uniformly random and the coefficients themselves are uniform random numbers in $[-1, 1]$. The measurement matrix Φ is chosen as eight consecutive rows of the DFT matrix of size 64. Figure 2 shows the normalized mean square error (MSE) of the reconstructed

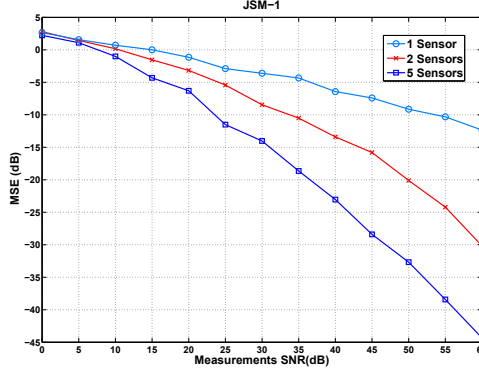


Fig. 2. JSM-1: Normalized MSE of the reconstructed signal vs. the measurements SNR for $n = 64$, $k_c = 4$, $m = 8$ and $N \in \{1, 2, 5\}$.

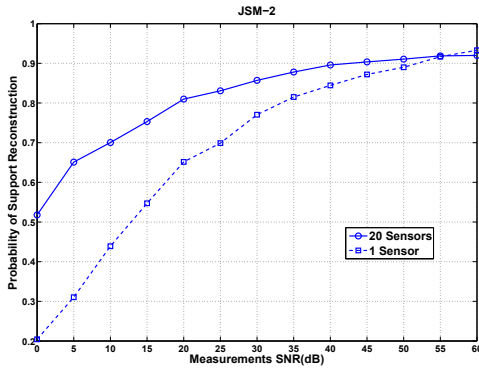


Fig. 3. JSM-2: Probability of correct support reconstruction vs. measurements SNR for $n = 128$, $k_c = 8$, $k_i = 3$, $m = 33$ and $N \in \{1, 20\}$.

signal vs. the SNR of the measurements for three different number of sensors, averaged over 500 trials. The nonzero coefficients are found by the least squares method. Note that using a few sensors results in a big improvement in the reconstruction performance, even if the number of measurements for each individual sensor is the minimum possible for perfect reconstruction in the noiseless case using our annihilating filter based recovery.

In the JSM-2 setup, the sparse signals are of length $n = 128$, the common support has dimension $k_c = 8$ and we let $k_i = 3$ for the innovation part. The number of measurements per sensor is fixed to $m = 33$. The simulation result is shown in Figure 3. For each point on the curves, we count the number of correct positions in 500 trials and divide it by the total number of positions. The non-zero coefficients for the common part are uniform random numbers in $[-1, 1]$ while we let the coefficients for the innovation part to be uniform random in $[-\alpha_{max}^c, \alpha_{max}^c]$ where α_{max}^c is the maximum absolute value of the coefficients in the common part of each signal. This is because we assume that the innovation part is not much stronger than the common part. The result shows the gain we get by the joint recovery of signals at the base station. One can see that for JSM-2 and at high SNR, the curve for the single sensor case is above the curve for multiple sensors. The reason is that for multiple sensors and high SNR values, the dominant noise in finding the common support set is the innovation part and not the white noise added to the measurements. This factor is not present in the single sensor case.

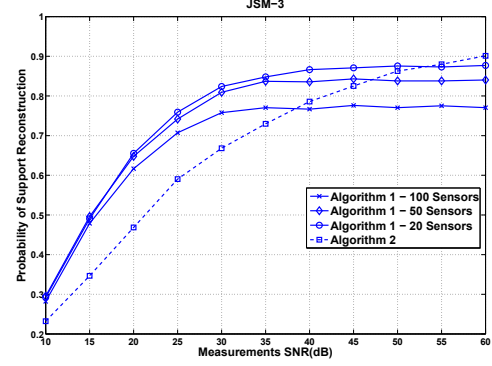


Fig. 4. JSM-3: Probability of correct support reconstruction vs. measurements SNR for $n = 64$, $k_i = 4$ and $m = 24$.

Finally consider the JSM-3 setup in which signals are of length $n = 64$. The common signal is of dimension 64 generated as a vector of uniform random numbers in $[-1, 1]$. The innovation part of the signals is of support length $k_i = 4$ and their coefficients are uniform in $[-1, 1]$. We choose number of measurements equal to $m = 24$. In Figure 4 we plot the probability of correct support reconstruction for different noise levels added to the measurements. Note that in algorithm 2, the number of sensors has no effect on the recovery performance because the innovation support is not common between the signals. If the support set is the same for the innovation parts, we are back to the JSM-1 scenario and the algorithm 2 can benefit from more sensors. Also note that there is a saturation effect in algorithm 1 because for high SNR, the dominant noise is the part coming from the common non-sparse part estimation and not from the white noise added to the measurements.

5. REFERENCES

- [1] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, pp. 1289 – 1306, April 2006.
- [2] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, pp. 489 – 509, February 2006.
- [3] A. Hormati and M. Vetterli, "Annihilating filter based decoding in the compressed sensing framework," *Wavelets XII*, August 2007.
- [4] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, pp. 1417–1428, June 2002.
- [5] M.F. Duarte, S. Sarvotham, D. Baron, M.B. Wakin, and R.G. Baraniuk, "Distributed compressed sensing of jointly sparse signals," *Conference record of thirty-ninth asilomar conference on signals, systems and computers*, vol. 24, pp. 1537 – 1541, October 2005.
- [6] T. Blu, P. L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse sampling of signal innovations: theory, algorithms and performance bounds," *Submitted to IEEE Signal Processing Magazine*, May 2007.
- [7] J. A. Cadzow, "Signal Enhancement-A composite property mapping algorithm," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 36, pp. 49–62, January 1988.